New Phytologist

Research

Global diversity and distribution of three necrotrophic effectors in *Phaeosphaeria nodorum* and related species

Megan C. McDonald[1], Richard P. Oliver[2], Timothy L. Friesen[3], Patrick C. Brunner[1] and Bruce A. McDonald[1]

[1]Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, CH-8092, Zurich, Switzerland; [2]Australian Centre for Necrotrophic Fungal Pathogens, Department of Environment and Agriculture, Curtin University, Perth, WA 6845, Australia; [3]United States Department of Agriculture-Agriculture Research Service, Northern Crop Science Laboratory, Cereal Crops Research Unit, Fargo, ND 58105, USA

Author for correspondence:
Megan C. McDonald
Tel: +61 261255561
Email: megan.mcdonald@anu.edu.au

Received: *6 February 2013*
Accepted: *22 February 2013*

New Phytologist (2013) **199:** 241–251
doi: 10.1111/nph.12257

Key words: co-evolution, fungal effectors, NB-LRR, necrotrophic effectors, *Phaeosphaeria nodorum*, plant pathogen, population genetics.

Summary

- Population genetic and phylogenetic studies have shown that *Phaeosphaeria nodorum* is a member of a species complex that probably shares its center of origin with wheat (*Triticum aestivum* and *Triticum durum*). We examined the evolutionary histories of three known necrotrophic effectors (NEs) produced by *P. nodorum* and compared them with neutral loci.
- We screened over 1000 individuals for the presence/absence of each effector and assigned each individual to a multi-effector genotype. Diversity at each NE locus was assessed by sequencing *c*. 200 individuals for each locus.
- We found significant differences in effector frequency among populations. We propose that these differences reflect the presence/absence of the corresponding susceptibility gene in wheat cultivars. The population harboring the highest sequence diversity was different for each effector locus and never coincided with populations harboring the highest diversity at neutral loci. Coalescent and phylogenetic analyses showed a discontinuous presence of all three NEs among nine closely related *Phaeosphaeria* species. Only two of the nine species were found to harbor NEs.
- We present evidence that the three described NEs of *P. nodorum* were transmitted to its sister species, *Phaeosphaeria avenaria tritici 1*, via interspecific hybridization.

Introduction

Necrotrophic fungal pathogens secrete a variety of necrotrophic effectors (NEs) (syn: host-specific/selective toxins) that interact in a gene-for-gene manner with host susceptibility genes (Oliver & Solomon, 2010). Typical plant defense responses involve the induction of pathogenicity-related genes, which leads to the production of antimicrobial compounds, the accumulation of reactive oxygen species and localized cell death. This defense response is known as the hypersensitive response (HR) (Jones & Dangl, 2006). A growing body of evidence supports the hypothesis that necrotrophic pathogens have taken advantage of the HR, using small secreted proteins or secondary metabolites to activate HR preceding fungal growth (Effertz *et al.*, 2002; Liu *et al.*, 2009, 2012; Lorang *et al.*, 2012). This group of molecules is collectively referred to as NEs.

Effectors are a class of pathogen proteins or metabolites whose function is to alter or suppress the host's normal immune response. Three NEs have been described for the fungal wheat pathogen *Phaeosphaeria nodorum* (Friesen *et al.*, 2006; Liu *et al.*, 2009, 2012). Each of these NEs is a small, secreted protein that displays a presence/absence polymorphism in natural field populations. *SnTox1* is a cysteine-rich protein that was shown to exhibit significant diversifying selection (Liu *et al.*, 2012). *SnTox3* has no known homology to any proteins available in public databases (Liu *et al.*, 2009). *SnToxA* has limited homology to a prokaryotic gene and also exhibits significant diversifying selection (Stukenbrock & McDonald, 2007). Transformation with any of the three NEs into a nonpathogenic fungal isolate was sufficient to induce necrosis on susceptible wheat (*Triticum aestivum*) cultivars (Friesen *et al.*, 2006; Liu *et al.*, 2009, 2012). Clamped homogeneous electric field (CHEF) gel analysis indicates that each NE is located on a different chromosome and each gene is located on a different scaffold in the genomic assembly (Hane *et al.*, 2007; Liu *et al.*, 2012).

While there is a growing list of shared properties associated with effectors (reviewed in Kamoun, 2007 and Stergiopoulos & de Wit, 2009), very little is known about the evolutionary origins of effector-encoding genes. Recent genome sequencing has revealed families of what have been termed core effector proteins within the Dothideomycetes (Stergiopoulos *et al.*, 2010, 2012). Identifying these core effectors relies on conservation of homologous elements within the effector proteins. The three characterized NEs of *P. nodorum* share no homology with any proteins for any fungal species available in GenBank and they do not appear to represent core effector proteins. Population genetic studies of effector loci have provided important insights into the evolutionary processes that affect NE loci within a species. A population genetic analysis

of the *NIP1* gene in *Rhynchosporium commune* showed that alteration of the effector protein sequence or deletion of the effector allele could lead to virulence (Schürch *et al.*, 2004). Studies on the obligate biotroph of flax (*Linum usitatissimum*), *Melampsora lini,* revealed high levels of nonsynonymous substitutions at the *AvrL567* locus and sectional insertions/deletions at the *AvrM* locus that alter or abolish recognition by their corresponding resistance (R) genes (Dodds *et al.*, 2006; Ellis *et al.*, 2007). Virulence alleles in *Leptosphaeria maculans* were attributed to both deletion of the *AvrLm6* locus and introduction of early stop codons by repeat-induced point mutation (RIP) (Fudal *et al.*, 2009; Van de Wouw *et al.*, 2010).

For NEs, deletion of the effector gene leads to loss of the virulent phenotype on hosts with compatible genetic backgrounds. To date, effector studies have focused on individual genes or small groups of effector genes in a small number of individuals (Dodds *et al.*, 2006; Barrett *et al.*, 2009; Liu *et al.*, 2009; Chuma *et al.*, 2011; Liu *et al.*, 2012), though it is clear that fungal populations are large and capable of harboring high levels of NE diversity. As more effector genes are discovered and characterized, a key question has become: how did this class of genes originate within fungal pathogens? Understanding the evolutionary origins of these genes could provide significant insights into the mechanisms involved in pathogen emergence and host specificity.

The horizontal transfer of *SnToxA* from *P. nodorum* to *Pyrenophora tritici-repentis* is thought to have led to the emergence of *P. tritici-repentis* as the tan spot pathogen on wheat (Friesen *et al.*, 2006). Detection of this horizontal gene transfer (HGT) event was made possible by the high sequence similarity between the *SnToxA* alleles found in *P. nodorum* and the *PtrToxA* allele found in *P. tritici-repentis.* It remains unknown if the NEs present in *P. nodorum* are the result of a long co-evolutionary process between the pathogen and its hosts, or alternatively if the NEs were acquired more recently via other mechanisms such as horizontal transfer or interspecific hybridization.

This study focuses on the population genetics and evolutionary history of *SnTox1, SnTox3* and *SnToxA* in *P. nodorum* and its closest known relatives. We assessed the global distribution and geographic diversity for all three effectors. We determined the presence or absence of each gene in over 1000 global isolates using both PCR and Southern hybridization. We calculated the frequency of each NE over spatial scales ranging from fields to continents and generated multi-effector genotypes to determine if selection was operating on the combination of NEs. To gain insight into the ancestral origin of these NEs in *P. nodorum*, we assessed the presence or absence of each NE in eight recently described sister species. Finally, we sequenced each NE in several hundred global strains to compare NE sequence diversity with previously published population genetic studies based on neutral markers.

Materials and Methods

Isolate collection

Isolates used in this study are described in Table 1. Fungal hyphae were transferred to 50 ml of yeast sucrose broth (YSB; 10 g l^{-1} yeast extract and 10 g l^{-1} sucrose) and grown on a rotary shaker for 3 d at 120 rpm at 18°C. Fungal biomass was lyophilized and ground into a powder and total DNA was extracted using the DNeasy Plant Mini DNA extraction kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's instructions. Extracted DNA was diluted 1 : 500 with sterile water for PCR amplification.

Data collection

PCR amplification was performed in 20-μl reactions containing 0.05 μM of each primer (Microsynth, Balgach, Switzerland), 1× Dream Taq Buffer (MBI Fermentas, Amherst, NY, USA), 0.4 μM dNTPs (MBI Fermentas) and 0.5 units of Dream Taq DNA polymerase (MBI Fermentas). The PCR cycle parameters were: 2 min of initial denaturation at 96°C followed by 35 cycles of 96°C for 30 s, annealing for 45 s and extension at 72°C for 1 min. A final 7-min extension was performed at 72°C. PCR products were purified to remove unincorporated nucleotides and primers using NucleoFast® 96 PCR plates (Macherey-Nagel, Oensingen, Switzerland). Details of the annealing temperature and primers used have been published previously; see Friesen *et al.* (2006) for SnToxA, Liu *et al.* (2009) for SnTox3 and Liu *et al.* (2012) for SnTox1.

Sequencing reactions were conducted in a 10-μl volume using the BigDye® Terminator v3.1 Sequencing Standard Kit (Life Technologies, Applied Biosystems, Grand Island, NY, USA) with both the forward and reverse primers. The cycling parameters were 96°C for 2 min followed by 55 or 99 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 4 min. PCR products were cleaned with the illustra™ Sephadex™ G-50 fine DNA Grade column (GE Healthcare, Pittsburgh, PA, USA) according to the manufacturer's recommendations and sequenced with a 3730xl Genetic Analyzer (Life Technologies, Applied Biosystems). Alignment of forward and reverse sequences was performed in SeqScape software V2.5 (Life Technologies, Applied Biosystems). Quality screening and ambiguous base calls were edited by hand using SeqScape. Final alignments were exported and re-aligned using ClustalW, implemented online using the Max Planck Institute of Bioinformatics Toolkit (http://toolkit.tuebingen.mpg.de). Sequences that have been published previously are SnToxA accession numbers DQ423483; EF108451–EF108463; SnTox3 accession number FJ823644; SnTox1 accession numbers JN791682–JN791693. New sequence haplotypes described in this manuscript are deposited in GenBank under accession numbers JX997397–JX997421.

Genomic DNA was blotted onto nylon membranes using a Bio-Dot microfiltration apparatus (Bio-Rad) following the instructions in the user manual. Radioactive probes were synthesized using purified PCR product (purification using NucleoFast® 96 PCR plates). Probe labeling with ^{32}P was performed with 25–50 ng μl^{-1} of template DNA using the NEBlot Kit (NE Biolabs, Ipswich, MA, USA) following the manufacturer's recommendations. Unincorporated nucleotides were filtered out using illustra NICK™ columns (GE Healthcare). DNA hybridization, membrane washing and image acquisition were performed as described previously (McDonald & Martinez, 1990).

Table 1 List of regional collections, sample sizes and host source for all isolates by species

Regions	Year	Collector	Location	*n*	Host source
Phaeosphaeria nodorum					
Iran	2005	R. Sommerhalder	Golestan Province	27	Wheat (ears)
	2010	M. Razavi	Golestan Province	13	Wheat (ears)
Central Asia	2005	R. Sommerhalder	Azerbaijan	7	Wheat
	2003–2004	H. Maraite, E. Duveiller	Kazakhstan	32	Wheat
	2004	H. Maraite, E. Duveiller	Tajikistan	7	Wheat
	2003	H. Maraite, E. Duveiller	Kyrgyzstan	1	Wheat
	2003	H. Maraite, E. Duveiller	Russia	9	Wheat/durum wheat
	2005	M. Abang	Syria	2	Wheat
China	2001	R. Wu	Fujian Province	115	Wheat
Europe	1992	M. Shaw	England	9	Wheat
	2005	E. Stukenbrock	Denmark	54	Triticale/wheat
	2005	E. Blixt	Sweden	55	Wheat
	1994	S. Keller	Switzerland	113	Wheat
USA	1991	G. Milus, M. Gray	Arkansas	94	Wheat (seed)
	1993	G. Shaner	Indiana	8	Wheat
	1993	L. Francl	North Dakota	12	Wheat
	1979	J. Krupinsky	North Dakota	2	*Agropyron* spp.
	1998	J. Krupinsky	North Dakota	9	Barley
	2003	P. Lips	OH03 Sn1501	1	Wheat
	2006	T. Friesen	North Dakota	6	Wheat
	1992	G. Bergstrom	New York	46	Wheat
	1993	M. Schmidt	Oregon	92	Wheat
	1992	B. McDonald	Texas	92	Wheat
South Africa	1994	P. Crous	Southwestern Cape	74	Wheat
	2007	Z. Pretorius	Western Cape	112	Wheat
Australia	2001	B. McDonald	Narrogin	73	Wheat
Total *P. nodorum*				1065	
Phaeosphaeria avenaria tritici 1					
Canada	1991	R. Clear	Alberta	65	Wheat (seed)
	1991	R. Clear	Manitoba	7	Wheat (seed)
	1991	R. Clear	Saskatchewan	36	Wheat (seed)
USA	1993	L. Francl	North Dakota	12	Wheat
	2005	T. Friesen	North Dakota	8	Wheat
Iran	2005	M. Razavi	Golestan Province	6	Wheat (ears)
	2010	M. Razavi	Golestan Province	18	Wheat (ears)
Total *Pat1*				152	

Data analysis

Contingency χ^2 tests and Fisher's exact tests Effector presence or absence was summarized as allele frequencies within each population. Contingency χ^2 tests were used to test for independence of effector frequencies between populations. The program CHIFISH was used to calculate both Fisher's exact test and Pearson's χ^2 test for each locus separately. *P*-values for independence based on the combined effector frequencies were calculated by summing the χ^2 test statistics and degrees of freedom, or by combining the *P*-values from Fisher's exact test using Fisher's method (Ryman & Jorde, 2001; Ryman, 2006). A Bonferroni correction was applied to correct *P*-values for multiple testing. After confirming the independence of each population, the expected number of multilocus genotypes was calculated. The frequency of each effector within a population was used as the expected probability of sampling an individual carrying the corresponding allele. For three bi-allelic loci there are eight possible multilocus genotypes (*ToxA*+*Tox3*+*Tox1*+; *ToxA*–*Tox3*–*Tox1*– etc.). The expected number of multilocus genotypes was compared with the observed number of multilocus genotypes using Fisher's exact test in each population (Fisher, 1925). This test was deemed to be most appropriate for large multinomial data sets with small expected values ($n < 5$) (Maiste & Weir, 2004). Fisher's exact tests were performed using the statistics package implemented in R. For contingency tables larger than 2×2, an estimation of the *P*-value was made using Monte Carlo simulations. Our test was conducted with MonteCarlo = TRUE with 10 000 replicates (Patefield, 1981). Each population was tested for independence separately, so no *P*-value corrections were applied.

Gene diversity and haplotype networks Effector sequences were collapsed into haplotypes using Map, implemented in the Java-based program package SNAP workbench (Price & Carbone,

2005; Aylor *et al.*, 2006). This program implements several individual programs into one platform in order to facilitate analysis of population parameters (Price & Carbone, 2005). The haplotype alignment was used in TCS 1.3 to generate the most parsimonious haplotype network (Clement *et al.*, 2000). This program utilizes statistical parsimony methods to infer unrooted cladograms based on Templeton's 95% parsimony connection limit. Mutational steps resulting in nonsynonymous changes were identified manually using SEQUENCHER v4.8 (Gene Codes Corp., Ann Arbor, MI, USA).

Rarefaction Rarefaction was used to identify regions of highest sequence diversity among populations with unequal sampling sizes. Rarefaction analysis was conducted using the method described by Simberloff (1978) and implemented in an online calculator (http://www.biology.ualberta.ca/jbrzusto/rarefact.php#Inputs). Rarified sample sizes from 10 to 20 were used to estimate the number of sequence alleles by global region. The output given is the rarefied allele count and standard deviation at each of the given sample sizes. Standard deviations were used to calculate the standard error (SE). The number of sequence alleles and the 95% confidence interval (1.96 × SE) were plotted for each sample size and population using R.

Results

Frequencies of NEs differ among populations but are randomly assorted within populations

Effector presence or absence was determined with gene-specific PCR primers for over 1000 isolates across seven major regions, including North America, Europe, Iran, Central Asia, China, South Africa and Australia (Table 1). Results from PCR assays were compared to Southern hybridization assays using 193 isolates for *SnToxA,* 284 isolates for *SnTox3* and 242 isolates for *SnTox1*. The average disagreement between the PCR and hybridization assays across all three loci was 4%. Sixteen out of the 30 discrepancies involved the highly polymorphic *SnTox1* locus and six of these were in the Central Asian population. When the two assays were not in agreement, the result from the hybridization assay was included in further analyses. Effector frequencies were calculated for each region and for the 16 subpopulations; variances were estimated using 100 bootstrapped replicates with replacement by region (Fig. 1). *SnToxA* was present at a lower global frequency than the other two effectors. *SnTox1* was found at a frequency > 60% in all continental regions. All three effectors were present at high frequencies in South Africa and Australia.

Effector frequencies differed significantly among field populations for all three loci according to contingency χ^2 tests and Fisher's exact tests (Supporting Information Table S1). In many cases, effector frequencies among field populations within a region (e.g. Arkansas and Texas within North America or Sweden and Denmark within Europe) also differed significantly.

The complete multilocus effector presence/absence genotype was determined for 945 strains. Given three effector loci, there are eight possible multilocus presence/absence genotypes. Based on the observed effector frequencies in each population, the expected number for each multilocus genotype under random mating was estimated and compared to the observed number of multilocus genotypes (Fig. 2). Fisher's exact test was used to determine if the number of observed genotypes differed from the number of expected genotypes. The only population with a significant deviation from expected proportions was Oregon (Table 2). Values for European and North American populations were summed to generate the graph shown in Fig. 2. Sample sizes for each NE locus are given along with the number of isolates for which complete genotypes were scored (Fig. 2).

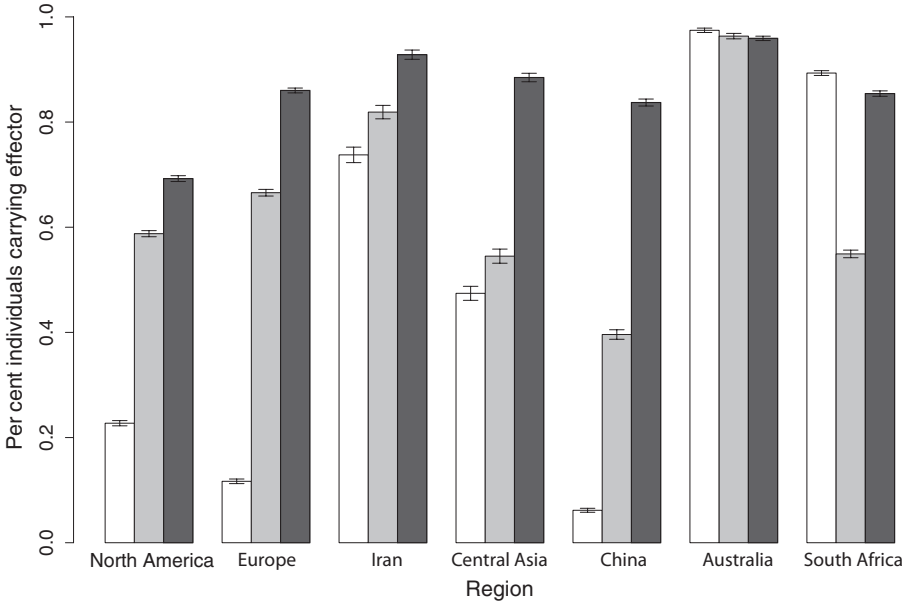


Fig. 1 Frequency of *Phaeosphaeria nodorum* necrotrophic effectors (NEs) is different between regions. Bar plots show the frequency of each NE within *P. nodorum* in seven global regions. Bootstrap re-sampling with replacement was performed on each data set to generate variance for the mean frequency of each NE. 95% confidence intervals are plotted for the bootstrap replicates.

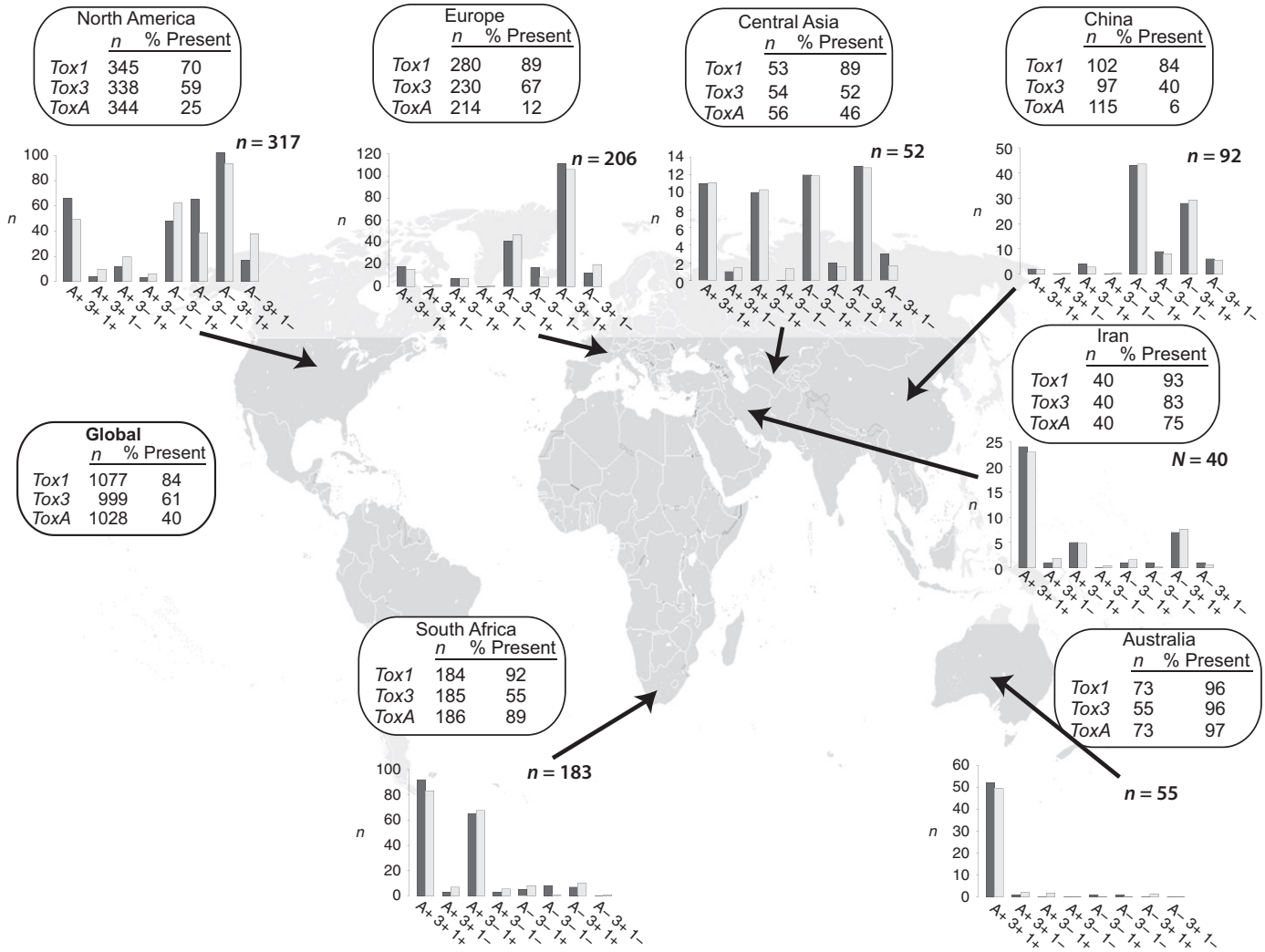


Fig. 2 The observed number of multi-effector genotypes matches the expected number under an assumption of random mating in *Phaeosphaeria nodorum*. The total number of *P. nodorum* isolates scored for the presence/absence of each locus is shown under N and alongside is the per cent presence of each locus in the sample. The bar plots show the observed multi-effector genotypes compared to the expected number of genotypes calculated from the effector presence/absence data. The total number of individuals for which all three necrotrophic effectors (NEs) were scored is listed as *n* = in the top right corner of the bar plot.

Effector sequence alleles are diverse and globally distributed

We sequenced *SnTox1, SnTox3* and *SnToxA* from 295, 203 and 178 individuals, respectively. The 295 *SnTox1* sequences collapsed into 18 nucleotide haplotypes, which encoded 14 different proteins (Fig. 3a). The *SnTox1* haplotype network contains three loops, indicating that some of the alleles originated through intragenic recombination. *SnTox3* collapsed into 11 nucleotide haplotypes encoding six different proteins (Fig. 3b). The 178 sequences from *SnToxA* collapsed into 17 unique haplotypes. Two *SnToxA* alleles with nonsense mutations were detected, one reported earlier in South Africa (Stukenbrock & McDonald, 2007) and a second one found in New York. Excluding these two nonsense haplotypes, there were nine different SnToxA proteins (Fig. 3c). Evidence for intragenic recombination was also found at the *SnToxA* locus.

Effector sequence diversity does not correspond with the center of diversity for neutral markers in *P. nodorum*

The region with the highest diversity for neutral markers in *P. nodorum* corresponds with the acknowledged origin of its wheat host in the ancient Fertile Crescent (Stukenbrock *et al.*, 2006; Balter, 2007; McDonald *et al.*, 2012). Allele diversity for each effector locus was compared among populations using rarefaction analysis across sample sizes ranging from 10 to 20 individuals. The rarefied number of alleles and 95% confidence intervals for each sample size are shown in Fig. 4(a–c). For *SnToxA* the South African population clearly contained the highest number of alleles. For *SnTox1* both Europe and North America had the highest number of alleles. The average number of alleles found within each population was much lower for *SnTox3,* with North America and Australia showing the highest number of alleles. While the highest diversity based on

Table 2 Fisher's exact tests for observed versus expected number of genotypes by population

Region	Population	*P*-value
Iran		1.00
Central Asia		1.00
China		1.00
Europe	England	1.00
	Denmark	0.10
	Sweden	0.53
	Switzerland	0.98
USA	Arkansas	0.73
	Indiana	1.00
	North Dakota	1.00
	New York	0.27
	Oregon	0.02*
	Texas	0.26
South Africa		0.13
Australia		0.43

*Number of observed genotypes deviates significantly from the number of expected genotypes.

neutral sequence loci and microsatellite loci was found in Iran (McDonald *et al.*, 2012), the highest diversity for each effector locus was always outside of Iran (Fig. 4). The number of private alleles (alleles found in only one region) is summarized in Fig. 4(d). Private sequence alleles comprised 56% (10 out of 18) of the *SnTox1* haplotypes, 64% (7 out of 11) of *SnTox3* haplotypes and 76% (13 out of 17) of *SnToxA* haplotypes.

Phylogenetic distribution of NEs among Phaeosphaeria spp

Phaeosphaeria nodorum is one of nine members of a species complex infecting wheat and other grasses (McDonald *et al.*, 2012). The presence or absence of the three NEs was determined in all nine species using both PCR and Southern hybridization. The three effectors were found only in *P. nodorum* and *Phaeosphaeria avenaria* f. sp. *tritici 1* (*Pat1*). Fig. 5 shows the coalescent multilocus species trees (adapted from McDonald *et al.*, 2012) for the nine *Phaeosphaeria* species as well as the frequency of each effector within *P. nodorum* and *Pat1*. Only one *Pat1* isolate carried *SnTox3* (Fig. 5).

The number of sequence alleles for each effector in both species is shown in Fig. 5. The 37 *SnTox1* sequences from *Pat1* collapsed into a single haplotype that was shared with both North American and Iranian *P. nodorum* isolates. The single *SnTox3* sequence found in *Pat1* was also the most frequent *P. nodorum* haplotype. Three *SnToxA* alleles were found among the 14 *Pat1* isolates sequenced. One *SnToxA* haplotype sequence was unique to *Pat1* but had the same amino acid sequence as the most frequent effector allele in the *SnToxA* network. The other two *Pat1* alleles were found in the two most common *P. nodorum* haplotypes. The overall diversity of *Pat1* sequence alleles was far lower than found in *P. nodorum*, with all but one of the alleles identical in sequence to a common *P. nodorum* allele. This pattern is consistent with the hypothesis that *Pat1* acquired all three NEs recently from *P. nodorum*.

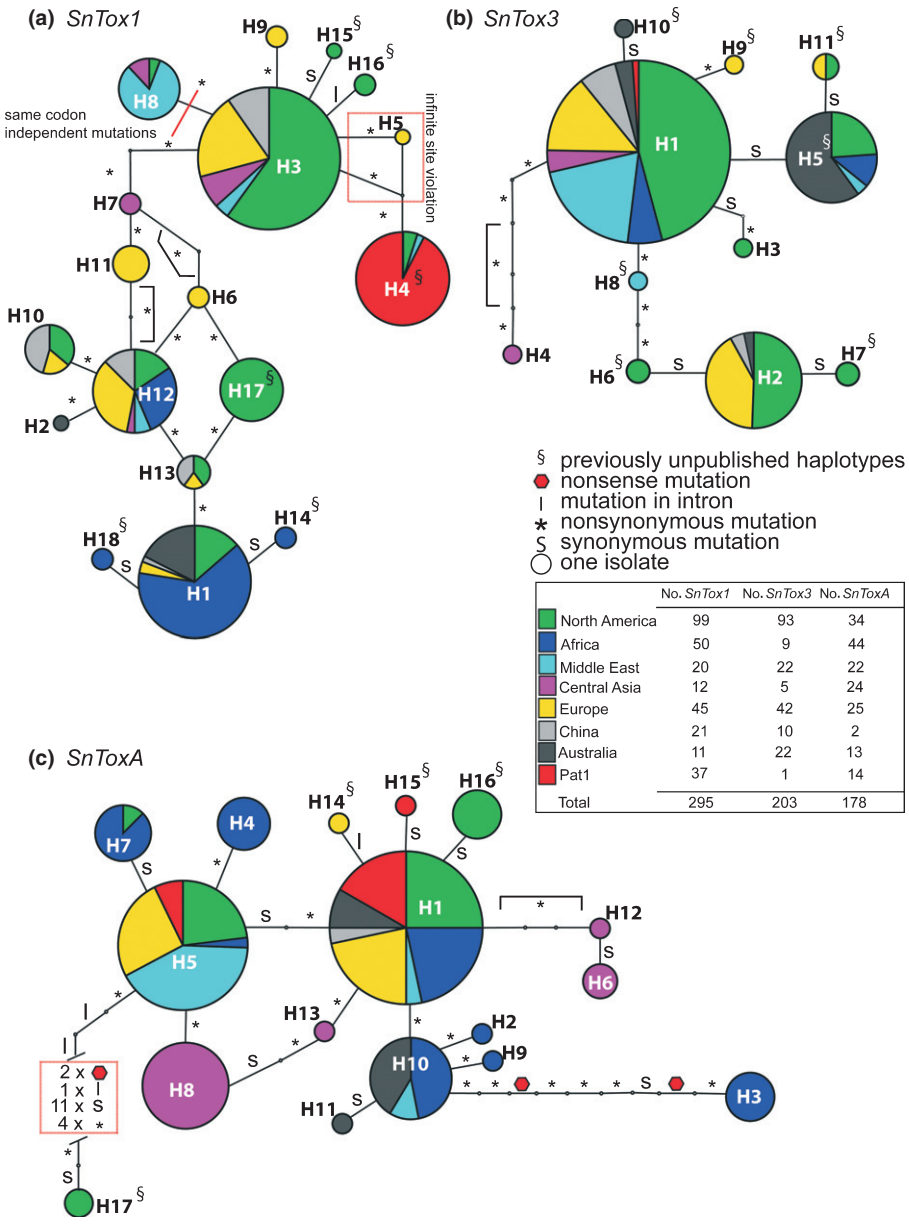
Discussion

Phylogenetic analysis of the *Phaeosphaeria* species complex revealed that only two out of nine closely related species carry the NEs *SnTox1*, *SnTox3* and *SnToxA*. All three NEs are diverse at the amino acid level with a high proportion of population-specific sequence alleles. Rarefaction analysis indicated that the center of diversity for each NE did not correspond with previous population genetic studies that identified the highest levels of diversity at neutral genetic loci in Iran (McDonald *et al.*, 2012).

Our analyses showed that the likelihood of an individual carrying one of the three NEs is population dependent. *Phaeosphaeria nodorum* is a sexual pathogen with a large effective population size that exhibits high levels of gene flow over continental scales (e.g. between Oregon, Texas and New York within North America; Keller *et al.*, 1997; Stukenbrock *et al.*, 2006). Based on gene flow estimates from neutral microsatellite loci, the expected frequencies of NEs in pathogen populations are not expected to differ among populations within a continent. Instead, we found significant differences in NE frequencies among many field populations that did not differ for neutral markers (Table S1). We believe that the differences in NE frequency among populations reflect differences in the frequencies of the corresponding host NE sensitivity genes among regions. As already shown for SnToxA and the corresponding host sensitivity protein Tsn1, we hypothesize that the activity of SnTox1 and SnTox3 depends upon an interaction with a corresponding host sensitivity protein which is present in some wheat cultivars but absent in others (Friesen *et al.*, 2006; Liu *et al.*, 2009, 2012). Unless there is a secondary virulence function, in the absence of a host sensitivity protein, pathogen strains carrying the corresponding NE have no fitness advantage and the NE gene is expected to be essentially neutral and subject to genetic drift (Tan *et al.*, 2012).

This highlights one of the main challenges associated with controlling globally disseminated pathogens with a high capacity for gene flow. Susceptible cultivars planted within the dissemination range of pathogen populations carrying an NE could rapidly select for pathogen populations carrying the NE. Thus, breeding efforts should be coordinated across large geographic regions to eliminate known susceptibility genes and reduce the frequencies of NEs at the continental scale. This type of effort is now underway in Australia to eliminate the *Tsn1* gene that encodes susceptibility to *SnToxA* and the sensitivity locus for Tox3 (Oliver & Solomon, 2010; Waters *et al.*, 2011).

A similar study that illustrated the dynamics of multiple effector loci in large natural populations was recently completed using the flax rust pathogen *Melampsora lini*. Thrall *et al.* (2012) found dramatic fluctuations in the frequency of *M. lini* avirulence alleles across multiple loci and they were able to correlate these fluctuations with the susceptibility of the corresponding host populations. Their analyses show how rapidly the genotype frequencies of host and pathogen can change in a gene-for-gene system experiencing antagonistic co-evolution. While we did not measure the sensitivity of the host in each field population, the large differences in local NE frequency despite high levels of neutral gene



	No. *SnTox1*	No. *SnTox3*	No. *SnToxA*
North America	99	93	34
Africa	50	9	44
Middle East	20	22	22
Central Asia	12	5	24
Europe	45	42	25
China	21	10	2
Australia	11	22	13
Pat1	37	1	14
Total	295	203	178

Fig. 3 Haplotype networks for each necrotrophic effector (NE) and sample sizes for sequence data by region. (a) The *SnTox1* haplotype network for the 18 sequence alleles found in *Phaeosphaeria nodorum* and *Phaeosphaeria avenaria tritici 1* (*Pat1*). *Pat1* haplotypes are drawn in red; all other haplotypes are *P. nodorum.* Circle sizes reflect the frequency of each haplotype and colors correspond to the global region where it was found. Haplotype numbers, denoted 'H#', correspond to previously published sequences. Newly described haplotypes are marked with §. Nonsynonymous mutations are marked with asterisks and synonymous mutations or mutations in the intron are labeled with 'S' and 'I', respectively. Open circles represent missing haplotypes in the network. Loops in the network indicate potential intra-locus recombination events. Unique features of the network are labeled with red lines or boxes and a short description. Black brackets indicate mutations that occurred within the same codon. (b) The *SnTox3* haplotype network. (c) The *SnToxA* haplotype network. The two nonsense haplotypes are labeled with a red hexagon.

flow suggest that there is very strong selection operating on NEs at the field level.

Despite significant differences in NE frequency between field populations, the distribution of multi-effector genotypes within all but one of the 16 field populations did not differ from the expectation under random mating of neutral markers (Table 2, Fig. 2). Traditionally, pathogens carrying particular combinations of avirulence or effector genes are classified into races (Barrett *et al.*, 2009), analogous to the multi-effector genotyping conducted in this study. The 'cost of carrying' is believed to drive the loss or alteration of the avirulence gene. Our finding of random associations among effector alleles within a population suggests that there is little fitness cost associated with carrying particular combinations of NEs. We hypothesize that the 'carrying cost' of these effector genes is low in the absence of the corresponding host sensitivity allele. This finding is also consistent with the hypothesis that host cultivar is the main determinant of NE frequency in these pathogen populations.

The haplotype networks presented in Fig. 3 show a prevalence of nonsynonymous mutations. It was reported previously that *SnToxA* and *SnTox1* are under significant positive selection (Friesen *et al.*, 2006; Liu *et al.*, 2012), while *SnTox3* did not show evidence of positive selection (Liu *et al.*, 2009). The detection of positive selection operating on protein-coding genes with unknown function has become a powerful tool for identifying effectors in fungal genomes (Ma & Guttman, 2008; Raffaele *et al.*, 2010; Stukenbrock *et al.*, 2011; Saunders *et al.*, 2012). It is often assumed that the higher rates of nonsynonymous substitutions exhibited by pathogen effector genes reflect diversifying selection favoring novel effector variants that are not recognized by plant R proteins. Dodds *et al.* (2006) performed experiments that supported this assumption in the gene-for-gene interactions

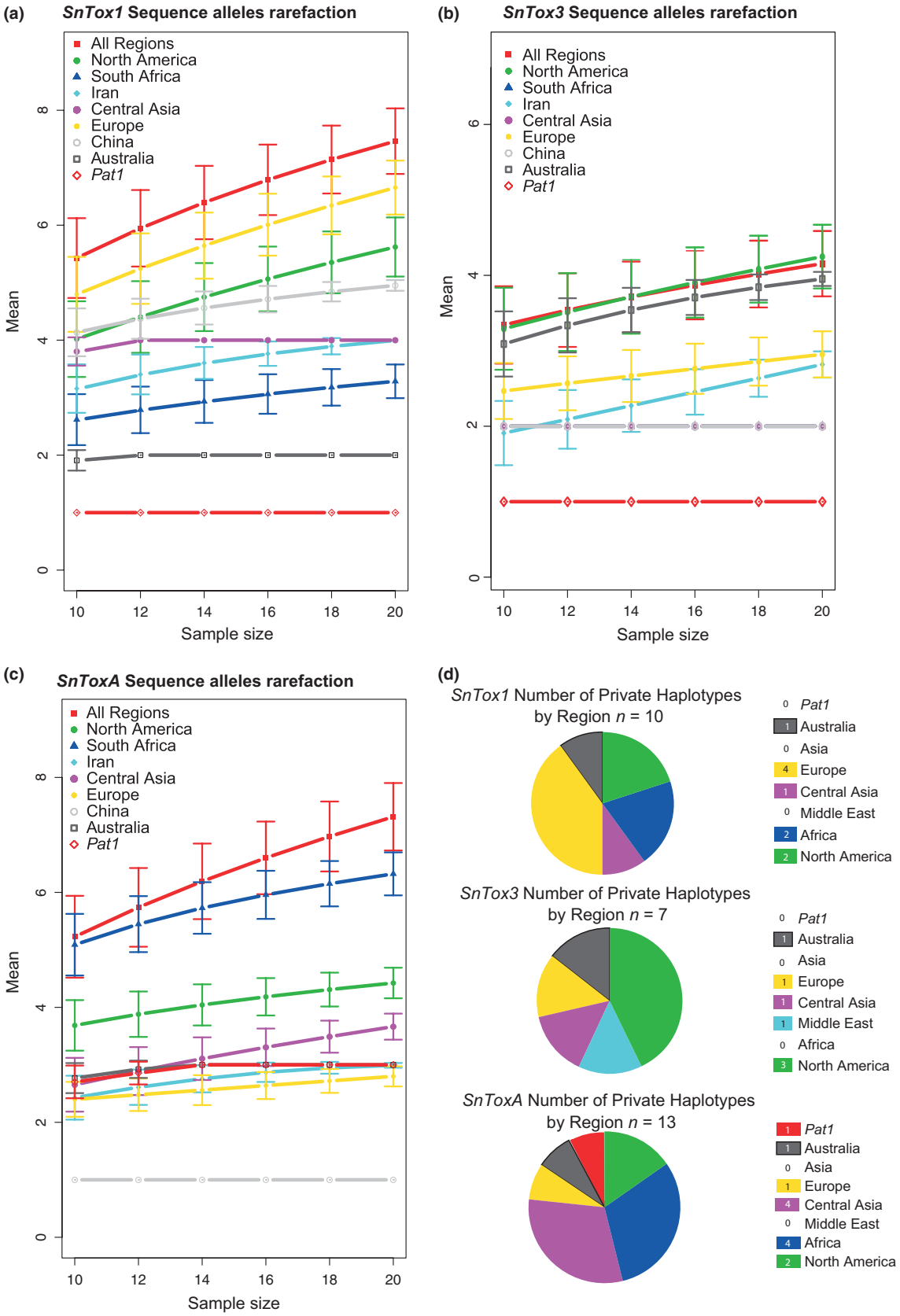


Fig. 4 The rarefied sequence allele count for each necrotrophic effector (NE) and number of private sequence alleles within each global region. (a–c) The mean number of alleles, after rarefaction, for sample sizes ranging from 10 to 20 randomly chosen individuals from each population. Vertical bars represent the 95% confidence interval for the mean number of alleles. All lines are *Phaeosphaeria nodorum* isolates, except for red lines with open diamonds, which are *Phaeosphaeria avenaria tritici 1*. (d) Pie charts showing the number and geographic location of private sequence alleles by NE locus.

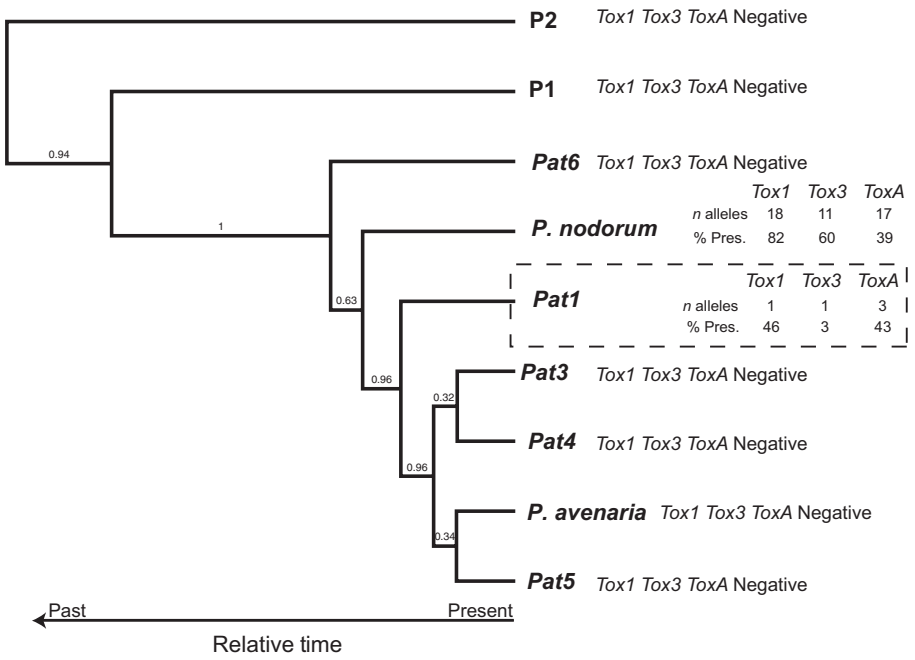


Fig. 5 Two out of nine *Phaeosphaeria* species carry necrotrophic effectors (NEs). The multilocus maximum-clade credibility tree adapted from McDonald *et al.* (2012) is shown. The tree is a coalescent tree with relative time on the *x*-axis and posterior probabilities for the branches shown. The two species within which NEs were found were *Phaeosphaeria nodorum* and *Pat1*. Other species are *Phaeosphaeria avenaria tritici 3*, *4*, *5 and 6*, *Phaeosphaeria avenaria* and unnamed species *Phaeosphaeria 1* and *2*, described in McDonald *et al.* (2012). The number of NE sequence alleles (*n* alleles) is shown above. The per cent NE presence within the species is shown below.

between *M. lini* and flax, where specific amino acid changes in Avr proteins altered or abolished recognition by their corresponding R proteins. For necrotrophic pathogens, it was hypothesized that the inverse process was operating, whereby higher rates of nonsynonymous substitution within NE proteins resulted from the pathogen tracking changes in the host susceptibility alleles (Stukenbrock & McDonald, 2007). An alternative hypothesis to explain the higher rates of nonsynonymous substitution seen in NEs is that positive selection has favored mutant NE alleles that increase pathogen fitness through a quantitative increase in virulence.

This alternative hypothesis is supported by recent experimental studies showing that the most frequent SnToxA protein variant is significantly more active against identical wheat *Tsn1* alleles hypothesized to increase pathogen fitness (Tan *et al.*, 2012). The haplotype networks of both *SnTox1* and *SnTox3* exhibit two or more frequent and widely distributed protein variants that differ at two or more amino acid positions. We hypothesize that the most common protein variants in these networks induce significantly more necrosis than the less common protein variants in the network. Experimental testing of this hypothesis is now underway.

There has been a rapid expansion of literature describing fungal effectors as small, secreted proteins or small metabolites that interact with the host to suppress or alter the immune response (reviewed by Kamoun, 2007; Hogenhout *et al.*, 2009; Stukenbrock & McDonald, 2009; de Wit *et al.*, 2009). As more effectors are identified and characterized, an important question has become: what are the evolutionary origins of these effectors and how did pathogens acquire them? For some filamentous plant pathogens, acquisition of effectors appears to be through the horizontal transfer of conditionally dispensable chromosomes (Hatta *et al.*, 2002; Oliver & Solomon, 2008; Akagi *et al.*, 2009; Ma *et al.*, 2010). Within the *Dothideomycetes,* families of functionally conserved effector genes analogous to *Ecp2* and *Avr4* in the tomato (*Lycopersicon esculentum*) pathogen *Cladosporium fulvum* have been identified (Stergiopolous *et al.*, 2010; Stergiopolous *et al.,* 2012). The SnTox1 protein shares local similarity with the chitin-binding domain found in Avr4, but otherwise the proteins appear to be unrelated (Liu *et al.*, 2012). Among more distantly related organisms, the rapid increase in genome sequences has led to the detection of HGT events involving single genes across kingdoms (Richards *et al.*, 2011). These HGT events were identified as a result of high homology among gene sequences. With the exception of SnToxA found in *P. tritici-repentis*, the NEs of *P. nodorum* do not show homology with any known proteins in other organisms, making it impossible to reconstruct the evolutionary history of these proteins. Though all three proteins require a host gene to induce symptoms and show similar expression profiles during infection, they are not homologous with each other. Because of the absence of homologous sequences outside of *P. nodorum*, we relied on diversity measurements from large population samples coupled with coalescent analyses of the *Phaeosphaeria* species complex to infer some aspects of the evolutionary history of these genes.

As shown in Fig. 5, the last highly supported (posterior probability = 1) recent common ancestor of *P. nodorum* is shared with six sister species. Among these seven species, *SnToxA, SnTox1* and *SnTox3* were found only in *P. nodorum* and *Pat1*. The sequences of the NE genes found in *Pat1* indicate that they were acquired directly from *P. nodorum* via hybridization, as discussed below. Taken together, our findings are consistent with at least two different evolutionary scenarios that could explain the origins of NEs within the *Phaeosphaeria* species complex. In one scenario, all three NEs were present in a common ancestor of the nine characterized *Phaeosphaeria* species and were then lost or became highly diverged in seven of these species. Under this scenario, strong selection operating on these genes generated high

sequence divergence and explains our inability to detect homologous genes in the other species. We used low-stringency hybridization conditions to test this hypothesis, but could not detect homologous sequences by hybridization. Another possibility is that a presence/absence polymorphism exists for each NE in all nine species and the isolates included in the analysis were missing all three NEs. Under a second scenario, all three NE genes in *P. nodorum* were acquired horizontally from an unknown donor or series of donors. The geographic distribution of diversity for each NE is consistent with three separate HGT events.

McDonald *et al.* (2012) provided evidence of hybridization between *P. nodorum* and *Pat1* based on sequence analysis of the *β-tubulin* locus. All of the effector sequence alleles found in *Pat1* were also found in *P. nodorum*, although the two species differed significantly for conserved housekeeping genes. For *SnTox3* and *SnToxA*, the shared alleles were also the most frequent alleles found in *P. nodorum*. This NE sequence data provide additional support for the hypothesis that hybridization occurred between *P. nodorum* and *Pat1*. Based on these findings we postulate that all three effectors in *Pat1* were acquired from *P. nodorum* via inter-specific hybridization, though the extent and nature of this hybridization require further investigation.

This population genetic study is only the second to compare neutral marker diversity with diversity at effector loci to infer the evolutionary history of effector genes. The earlier population study on the barley (*Hordeum vulgare*) scald pathogen *Rhynchosporium commune* found that the highest diversity for neutral marker loci, including DNA sequences, restriction fragment length polymorphisms and microsatellites, corresponded with the highest diversity for the *NIP1* effector gene in Scandinavia (Schürch *et al.*, 2004; Brunner *et al.*, 2007), indicating that *NIP1* shared the same evolutionary history, and probably the same common ancestor, as the other *R. commune* genes. In contrast, the lack of geographic correspondence between toxin diversity and neutral marker diversity provides strong evidence that the NEs of *P. nodorum* do not share the same evolutionary history as the neutral loci.

The geographic region that harbored the highest sequence diversity was different for each effector gene. The highest diversity for *SnToxA* was found in South Africa, for *SnTox3* in North America and Australia, and for *SnTox1* in Europe. The finding of higher effector diversity in 'New World' populations, where wheat cultivation began only during the last few hundred years following the arrival of European colonists (South Africa *c.* 350 yr ago, Australia *c.* 200 yr ago and North America *c.* 500 yr ago), suggests that each of these effectors may be under strong regional selection or alternatively may have different geographic origins representing three separate HGT events. None of the populations harboring the highest effector diversity overlapped with the hypothesized center of origin of *P. nodorum* in the ancient Fertile Crescent. The original source of the three effector genes remains unknown, but the SnToxA protein has a domain consistent with a prokaryotic origin (Friesen *et al.*, 2006). McDonald *et al.* (2012) presented evidence that *P. nodorum* existed as a species before the domestication of wheat, probably as a fungal endophyte on grasses. We propose that these NE genes, whether inherited from an unknown common ancestor or acquired horizontally, enabled *P. nodorum* to emerge from this species complex as the dominant, specialized pathogen on wheat. Hybridization with its sister species *Pat1* enabled the transfer of these genes to a new fungal species, resulting in the emergence of another damaging, though closely related, pathogen.

Acknowledgements

The authors acknowledge the Genetic Diversity Center at ETH Zurich for providing sequencing facilities and the ETH Zurich for funding this research. We also thank the Australian Grains Research and Development Corporation and the USDA-NIFA AFRI Microbial Biology Program–Competitive grant no. 2009-04265 for additional funding.

References

Akagi Y, Akamatsu H, Otani H, Kodama M. 2009. Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. *Eukaryotic Cell* **8**: 1732–1738.

Aylor DL, Price EW, Carbone I. 2006. SNAP: Combine and Map modules for multilocus population genetic analysis. *Bioinformatics* **22**: 1399–1401.

Balter M. 2007. Seeking agriculture's ancient roots. *Science* **316**: 1830–1835.

Barrett LG, Thrall PH, Dodds PN, van der Merwe M, Linde CC, Lawrence GJ, Burdon JJ. 2009. Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. *Molecular Biology and Evolution* **26**: 2499–2513.

Brunner PC, Schürch S, McDonald BA. 2007. The origin and colonization history of the barley scald pathogen *Rhynchosporium secalis*. *Journal of Evolutionary Biology* **20**: 1311–1321.

Chuma I, Isobe C, Hotta Y, Ibaragi K, Futamata N, Kusaba M, Yoshida K, Terauchi R, Fujita Y, Nakayashiki H *et al.* 2011. Multiple translocation of the *AVR-Pita* effector gene among chromosomes of the rice blast fungus *Magnaporthe oryzae* and related species. *PLoS Pathogens* 7: e1002147.

Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.

Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang C-IA, Ayliffe MA, Kobe B, Ellis JG. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences, USA* **103**: 8888–8893.

Effertz RJ, Meinhardt SW, Anderson JA, Jordahl JG, Francl LJ. 2002. Identification of a chlorosis-inducing toxin from *Pyrenophora tritici-repentis* and the chromosomal location of an insensitivity locus in wheat. *Phytopathology* **2002**: 527–533.

Ellis JG, Dodds PN, Lawrence GJ. 2007. Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology* **45**: 289–306.

Fisher RA. 1925. *Statistical methods for research workers.* Edinburgh, UK: Oliver and Boyd.

Friesen TL, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, Faris JD, Rasmussen JB, Solomon PS, McDonald BA, Oliver RP. 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature Genetics* **38**: 953–956.

Fudal I, Ross S, Brun H, Besnard A-L, Ermel M, Kuhn M-L, Balesdent M-H, Rouxel T. 2009. Repeat-induced point mutation (RIP) as an alternative mechanism of evolution toward virulence in *Leptosphaeria maculans*. *Molecular Plant-Microbe Interactions* **22**: 932–941.

Hane JK, Lowe RG, Solomon PS, Tan KC, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE *et al.* 2007. Dothideomycete plant interactions illuminated by genome sequencing and EST analysis of the wheat. pathogen *Stagonospora nodorum*. *Plant Cell* **19**: 3347–3368.

Hatta R, Ito K, Hosaki Y, Tanaka T, Tanaka A, Yamamoto M, Akimitsu K, Tsuge T. 2002. A conditionally dispensable chromosome controls host-specific pathogenicity in the fungal plant pathogen *Alternaria alternata*. *Genetics* **161**: 59–70.

Hogenhout SA, Van der Hoorn RAL, Terauchi R, Kamoun S. 2009. Emerging concepts in effector biology of plant-associated organisms. *Molecular Plant Microbe Interactions* **22**: 115–122.

Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* **444**: 323–329.

Kamoun S. 2007. Groovy times: filamentous pathogen effectors revealed. *Current Opinion In Plant Biology* **10**: 358–365.

Keller SM, McDermott JM, Pettway RE, Wolfe MS, McDonald BA. 1997. Gene flow and sexual reproduction in the wheat glume blotch pathogen *Phaeosphaeria nodorum* (anamorph: *Stagonospora nodorum*). *Phytopathology* **87**: 353–358.

Liu Z, Faris JD, Oliver RP, Tan K-C, Solomon PS, McDonald MC, McDonald BA, Nunez A, Lu S, Rasmussen JB *et al.* 2009. SnTox3 acts in effector triggered susceptibility to induce disease on wheat carrying the *Snn3* gene. *PLoS Pathogens* **5**: e1000581.

Liu Z, Zhang Z, Faris JD, Oliver RP, Syme R, McDonald MC, McDonald BA, Solomon PS, Lu S, Shelver WL *et al.* 2012. The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. *PLoS Pathogens* **8**: e1002467.

Lorang J, Kidarsa T, Bradford CS, Gilbert B, Curtis M, Tzeng SC, Maier CS, Wolpert TJ. 2012. Tricking the guard: exploiting plant defense for disease susceptibility. *Science* **338**: 659–662.

Ma W, Guttman DS. 2008. Evolution of prokaryotic and eukaryotic virulence effectors. *Current Opinion In Plant Biology* **11**: 412–419.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B *et al.* 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**: 367–373.

Maiste P, Weir B. 2004. Optimal testing strategies for large, sparse multinomial models. *Computational Statistics & Data Analysis* **46**: 605–620.

McDonald BA, Martinez JP. 1990. Restriction Fragment Length Polymorphisms in *Septoria tritici* occur at a high frequency. *Current Genetics* **17**: 133–138.

McDonald MC, Razavi M, Friesen TL, Brunner PC, McDonald BA. 2012. Phylogenetic and population genetic analyses of *Phaeosphaeria nodorum* and its close relatives indicate cryptic species and an origin in the Fertile Crescent. *Fungal Genetics and Biology* **49**: 882–895.

Oliver RP, Solomon PS. 2008. Recent fungal diseases of crop plants: is lateral gene transfer a common theme? *Molecular Plant-Microbe Interactions* **21**: 287–293.

Oliver RP, Solomon PS. 2010. New developments in pathogenicity and virulence of necrotrophs. *Current Opinion In Plant Biology* **13**: 415–419.

Patefield W. 1981. Algorithm AS 159: an efficient method of generating random R×C tables with given row and column totals. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **30**: 91–97.

Price EW, Carbone I. 2005. SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics* **21**: 402–404.

Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, Thines M, Jiang RHY, Zody MC, Kunjeti SG, Donofrio NM *et al.* 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science* **330**: 1540–1543.

Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, Foster PG, Hall N, Talbot NJ. 2011. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proceedings of the National Academy of Sciences, USA* **108**: 15258–15263.

Ryman N. 2006. chifish: a computer program testing for genetic heterogeneity at multiple loci using chi-square and Fisher's exact test. *Molecular Ecology Notes* **6**: 285–287.

Ryman N, Jorde P. 2001. Statistical power when testing for genetic differentiation. *Molecular Ecology* **10**: 2361–2373.

Saunders DGO, Win J, Cano LM, Szabo LJ, Kamoun S, Raffaele S. 2012. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PLoS ONE* **7**: e29847.

Schürch S, Linde C, Knogge W, Jackson L, McDonald B. 2004. Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene *NIP1*. *Molecular Plant-Microbe Interactions* **17**: 1114–1125.

Simberloff D. 1978. Use of rarefaction and related methods in ecology. In: Dickson KL, Cairns J Jr, Livingston RJ, eds. *Biological data in water pollution assessment: quantitative and statistical analyses*. Easton MD, USA: American Society for Testing and Materials, 150–165.

Stergiopoulos I, de Wit PJGM. 2009. Fungal effector proteins. *Annual Review of Phytopathology* **47**: 233–263.

Stergiopoulos I, Kourmpetis YAI, Slot JC, Bakker FT, de Wit PJGM, Rokas A. 2012. In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. *Molecular Biology and Evolution* **29**: 3371–3384.

Stergiopoulos I, van den Burg HA, Ökmen B, Beenen HG, van Liere S, Kema GHJ, de Wit PJGM. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. *Proceedings of the National Academy of Sciences, USA* **107**: 7610–7615.

Stukenbrock EH, Banke S, McDonald BA. 2006. Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. *Molecular Ecology* **15**: 2895–2904.

Stukenbrock EH, Bataillon T, Dutheil JY, Hansen TT, Li R, Zala M, McDonald BA, Wang J, Schierup MH. 2011. The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. *Genome Research* **21**: 2157–2166.

Stukenbrock EH, McDonald BA. 2007. Geographical variation and positive diversifying selection in the host-specific toxin *SnToxA*. *Molecular Plant Pathology* **8**: 321–332.

Stukenbrock EH, McDonald BA. 2009. Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions. *Molecular Plant-Microbe Interactions* **22**: 371–380.

Tan K-C, Ferguson-Hunt M, Rybak K, Waters ODC, Stanley WA, Bond CS, Stukenbrock EH, Friesen TL, Faris JD, McDonald BA *et al.* 2012. Quantitative variation in effector activity of ToxA isoforms from *Stagonospora nodorum* and *Pyrenophora tritici-repentis*. *Molecular Plant-Microbe Interactions* **25**: 515–522.

Thrall PH, Laine A-L, Ravensdale M, Nemri A, Dodds PN, Barrett LG, Burdon JJ. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters* **15**: 425–435.

Van de Wouw AP, Cozijnsen AJ, Hane JK, Brunner PC, McDonald BA, Oliver RP, Howlett BJ. 2010. Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathogens* **6**: e1001180.

Waters ODC, Lichtenzveig J, Rybak K, Friesen TL, Oliver RP. 2011. Prevalence and importance of sensitivity to the *Stagonospora nodorum* necrotrophic effector SnTox3 in current Western Australian wheat cultivars. *Crop and Pasture Science* **62**: 556.

de Wit PJGM, Mehrabi R, Van den Burg HA, Stergiopoulos I. 2009. Fungal effector proteins: past, present and future. *Molecular Plant Pathology* **10**: 735–747.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Contingency Fisher's exact and Pearson's χ^2 tests for the frequency of *SnTox1*, *SnTox3* and *SnToxA* within North America, Europe and Asia

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.